Final Progress Report 1999

Grant Title:	Shape and Motion Categorization for Content-Based Image and Video Database Search
Grant Number:	N00014-96-1-0661
Institution:	Boston University
PI:	Stan Sclaroff
Phone:	(617) 353-8928
FAX:	(617) 353-6457
E-mail:	sclaroff@bu.edu
WWW:	http://www.cs.bu.edu/faculty/sclaroff/
Reporting Period:	June 1, 1996 -- Aug 31, 1999
Team Members:	Dr. Stan Sclaroff, Assistant Professor, partially supported Marco La Cascia, visiting research associate, partially supported John Isidoro, grad. research assistant, partially supported Lifeng Liu, grad. research assistant, partially supported Jonathan Alon, grad. research assistant, fully supported Saratendu Sethi, grad. research assistant, partially supported

Web pages and technical reports describing research in Boston University's Image and Video Computing Group can be found at http://www.cs.bu.edu/groups/ivc/Home.html

1. Project Summary

The aim of this project is to represent shape categories for interactive, image database search. The work involves developing representations of images of objects to facilitate content-based retrieval and indexing, especially of deformable objects.

Approach

Rather than directly comparing a candidate shape with all shapes in the database, we are developing methods that describe shape categories in terms of their relationship to a few deformable shape prototypes. This approach is related to the computer graphics technique of image metamorphosis (morphing). In morphing, a wide variety of in-between or novel views can be generated as warps between a relatively small collection of example views. This suggests an important way to obtain a low-dimensional, parametric description of shape: interpolate between known, *deformable prototype views*. For instance, given views of the extremes of a motion we can describe the intermediate views as a smooth combination of the extremal views.

The prototypes define a polytope in the space of the (unknown) underlying physical system's parameters. By measuring the amount of deformation between a new shape and prototype views, we locate the new shape in the coordinate system defined by the polytope. This coordinate in prototype space can be used for database indexing and fast search, and for motion tracking and categorization. Such a representation could also prove useful surveillance, low bit-rate video compression, target recognition and tracking, and medical image analysis.

This research is built on top of an existing shape representation framework called modal models. The underlying representation employs *modal models*, a deformable shape decomposition that allows users to specify a few example shapes and has the computer efficiently sort the set of objects based on the similarity of their shape. If desired, shapes can be more closely compared in terms of the types of nonrigid deformations (differences) that relate them to a few prototype shapes. Furthermore, the original shape can be reconstructed in terms of a linear combination of deformed basis images; thus, a semantics-preserving shape representation is obtained.

2. Summary of Progress

Aug 1998 -- Aug 1999

Continued experimental evaluation and improvement of nonrigid shape model software (active blobs). Also completed development and extensive testing of 2D active blobs embedded in the surface of a 3D rigid model (demo application: head tracking). The two resulting software implementations are available for public FTP.

Continued development of deformable model detection/segmentation algorithms. Focussed on the development of approximation algorithms to gain better performance in testing model hypotheses and region groupings. Tested in image database of fish.

Began development of automatic methods for prototype registration, selection, and category representation via the active blobs formulation. However, the challenges presented by efficient selection of prototypes in large databases remain to be addressed.

Investigated the use of active blobs for estimating 3D relative structure (shape) and motion given objects undergoing nonrigid deformation. Deforming surfaces are approximated as piece-wise planar and piece-wise rigid. This work may lead to new methods for estimating the 3D structure of objects for use in new types of content-based image and video database queries.

June, 1996 -- Aug 1998

Tested category representation methods for stability to noise. Developed and conducted experimental evaluation of shape-based comparison metrics for use in content-based retrieval.

Extended deformable shape framework to include intensity information. This will lead to new methods for content-based indexing and retrieval on shape *plus* appearance (pixel-wise match).

Developed algorithms for tracking and indexing deforming shapes in image plane (2D). This will help in the area of moving shape representation and motion-based indexing.

Developed extension of 2D tracker embedded in the texture map of a 3D rigid surface. This should allow better tracking of convex, translation/rotating 3-D objects. This approach led to a novel approach for 3D head tracking that can effectively model large head rotation and facial deformation.

Developed new methods for automatic detection and segmentation of shapes from images. Once trained, the system automatically segments deformed shapes from the background, while not merging them with similarly colored adjacent objects. The method can serve as a front end to shape-based image retrieval, and was tested in preliminary experiments on databases of 100-200 images.

3. Technical Transitions

Worked with Dr. Stephen Bogen, Boston University Medical School. The focus is on deformable shape segmentation/detection methods in finding certain types of cells in color microscope images of human blood. The goal is to develop automatic image search tools that can find macrophages in human blood samples. This could lead to earlier detection of certain cancers.

Shared software developed for this project with Sven Dickinson and Stan Dunn at Rutgers. The software is used as part of a project to develop new methods for content-based organization and search of digital image databases of dental X-rays.

Shared software developed for this project with Alex Pentland at the MIT Media Lab. Application of results from this project are planned in the area of deformable shape modeling algorithms for locating and tracking people in dynamic environments (for ONR MURI grant with U. of Maryland).

The modal framework is being used by researchers in Italy and the United Kingdom. Dell'Acqua, Gamba (U. of Pavia, Italy), and Mecocci (U. of Sienna) reported the use of modal matching for visual search in image databases using user sketches. Mike Syn (Oxford), extended the modal matching to 3D for use in biomedical data set analysis.

In collaboration with Ron Kikinis at Brigham and Women's Hospital, work was conducted to transfer deformable shape methods to biomedical applications. The focus was on developing 3-D shape models for tracking and anatomical structures in medical volume data for computer-assisted diagnosis and surgical planning.

4. Significant accomplishments

Tested deformable shape category representation sensitivity to noise. Tested in shape-based image database retrieval experiments. In the experiments, the formulation performed significantly better than the moment invariants technique. Journal paper describing the work appeared in *Pattern Recognition* 30(4), 1997.

Active blobs, a new region-based formulation for tracking 2D deforming shapes was developed, tested, and packaged for public distribution via FTP. This extended our framework to allow a combined representation for shape and appearance (texture). The method utilizes a robust estimation formulation that allows stable tracking in spite of noise, minor occlusion, and shadows. The models can robustly track nonrigidly deforming shapes at speeds approaching video frame-rate on a standard graphics workstation. A software implementation was made available in fall 1998: http://www.cs.bu.edu/groups/ivc/ActiveBlobs/. A number of researchers at other universities now have copies of the software. A journal manuscript describing the work has been submitted to IEEE T-PAMI.

New technique for deformable shape tracking that imbeds the 2-D nonrigid shape models in a rigid 3-D surface was developed. The approach's effectiveness has been demonstrated in a 3D head tracking and facial expression tracking application. A software implementation was made available in spring 1999: http://www.cs.bu.edu/groups/ivc/HeadTracking/. A journal manuscript describing the work has been submitted to IEEE T-PAMI.

New technique for segmentation and detection of deformable shapes in color images was developed. Statistical shape models are used to enforce the prior probablities on global, parametric deformations for each object class. Once trained, the system autonomously segments deformed shapes from the background, while not merging with adjacent objects or shadows. The formulation can be used to group image regions based on any image homogeneity predicate; e.g., texture, color, or motion. The recovered shape models can be used directly in content-based indexing of image databases. The approach's effectiveness has been demonstrated in a 2D shape-based search application. A journal manuscript describing the work has been submitted to IEEE T-PAMI.

5. Publications resulting from work done on this grant

Aug 1998 -- Aug 1999

1. Liu, L., and Sclaroff, S., "Automatic Deformable Shape Segmentation for Image Database Search Applications," *Proc. International Conf. on Visual Information*, Lecture notes in Computer Science 1614, pp. 601--608, Springer-Verlag, Jun., 1999.

2. La Cascia, M., and Sclaroff, S., "Fast, Reliable Head Tracking under Varying Illumination," *IEEE Conf. on Comp. Vision and Pattern Rec., (CVPR),* Vol. 1, pp. 604--610, Jun., 1999.

3. Liu, L., and Sclaroff, S., "Deformable Shape Detection and Description via Model-Based Region Grouping," *IEEE Conf. on Comp. Vision and Pattern Rec. (CVPR),* Vol. 2, pp. 21--27, Jun., 1999.

4. Sethi, S., and Sclaroff, S., "Combinations of Deformable Shape Prototypes," *CS TR-99-007*, Boston University, Jul., 1999.

5. Sclaroff, S., and Alon, J., "Non-Rigid Shape from Image Streams," *CS TR-99-006*, Boston University, Jul., 1999.

6. Sethi, S., and Sclaroff, S., "Combinations of Nonrigid Deformable Appearance Models," *SPIE Conf. on Telemanipulator and Telepresence Tech. VI, (SPIE 3840),* Sep., 1999 (to appear).

7. Alon, J., and Sclaroff, S., "Recovery of Piece-wise Planar and Piece-wise Rigid Models from Nonrigid Motion," *SPIE Conf.on Three-Dimensional Imaging, Optical Metrology, and Inspection V,* (SPIE 3835), Sep., 1999 (to appear).

8. Sclaroff, S. and Isidoro, J. "Active Blobs: Deformable Color Appearance Models," *IEEE Trans. on Pattern Analysis and Mach. Intelligence* (PAMI), (submitted).

9. Sclaroff, S., and Liu, L., "Deformable Shape Detection and Description via Model-Based Region Grouping," *IEEE Trans. Pattern Analysis and Mach. Intelligence (PAMI),* (submitted).

10. La Cascia, M., Sclaroff, S., and Athitsos, V., "Fast, Reliable Head Tracking under Varying Illumination: An Approach Based on Robust Registration of Texture-Mapped 3D Models," *IEEE Trans. Pattern Analysis and Mach. Intelligence (PAMI),* (submitted).

June, 1996 -- Aug 1998

1. Sclaroff, S., "Encoding Deformable Shape and Motion Categories for Efficient Content-Based Search," *Proc. First International Workshop on Image Databases and Multimedia Search,* pp. 107--114, Aug., 1996.

2. Zhang, W., Dickinson, S., Sclaroff, S., Marsic, I., Hawkins, S., Feldman, J., Dunn, S., "Searching Medical Image Databases by Image Content," in *Proc. Ninth Image and Multidimensional Signal Processing Workshop*, pp.146--147, Mar., 1996.

3. Sclaroff, S., "Deformable Prototypes for Encoding Shape Categories in Image Databases," *Pattern Recognition*, 30(4):627--642, Apr., 1997.

4. Liu, L. and Sclaroff, S., "Color Region Grouping and Shape Recognition with Deformable Models," *CS TR-97-019*, Boston University, Nov., 1997.

5. Sclaroff, S., "Distance to Deformable Prototypes: Encoding Shape Categories for Efficient Search," chapter in *Image Databases and Multi-Media Search,* A.W.M. Smeulders and R. Jain, Ed., pp. 149--164, World Scientific, Singapore, 1998.

6. Sclaroff, S. and Isidoro, J. "Active Blobs," *Proc. International Conf. on Comp. Vision* (ICCV), pp. 1146--1153, Jan. 1998.
7. Isidoro, J. and Sclaroff, S. "Active Voodoo Dolls: A Vision Based Input Device for Non-rigid Control", *Proc. Computer Animation 98,* Jun., 1998.
8. La Cascia, M., Isidoro, J, and Sclaroff, S., "Head Tracking via Robust Registration in Texture Map Images," *IEEE Conf. on Computer Vision and Pattern Recognition,* (CVPR) Jun., 1998.

6. Other publications

Aug 1998 -- Aug 1999

1. Rosales, R., and Sclaroff, S., "3D Trajectory Recovery for Tracking Multiple Objects and Trajectory Guided Recognition of Actions," *IEEE Conf. on Comp. Vision and Pattern Rec., (CVPR),* Vol. 2, pp. 117-121, Jun., 1999.
2. Sclaroff, S., La Cascia, M., Taycher, L., and Sethi, S., "Combining Textual and Visual Cues for Improved Performance in Image Search on the World Wide Web," *Computer Graphics and Image Processing: Image Understanding (CVGIP:IU),* Jul., 1999.
3. Rosales, R., and Sclaroff, S., "3D Tracking of Humans in Action and the Recognition of Human Body Configuration for Visually-Guided Interfaces," *SPIE Conf. on Telemanipulator and Telepresence Technologies VI, (SPIE 3840),* Boston, MA, Sep., 1999 (to appear).
4. Sclaroff, S., La Cascia, M., Sethi, S., and Taycher, L., "ImageRover Architecture," chapter in Multimedia Internet Information Retrieval, A. Picariello, editor (to appear).
5. Rosales, R., and Sclaroff, S., "Trajectory Guided Tracking and Recognition of Actions," *IEEE Trans. Pattern Analysis and Mach. Intelligence (PAMI)* (submitted).

June, 1996 -- Aug 1998

1. Pentland, A., Picard, R., and Sclaroff, S., "Photobook: Tools for Content-Based Manipulation of Image Databases," *International Journal of Computer Vision (IJCV),* 18(3):233--254, Jun., 1996.
2. Pentland, A., Picard, R., and Sclaroff, S., "Photobook: Content-Based Manipulation of Image Databases," chapter in *Multimedia Tools and Applications*, B. Furht, Ed., pp. 43--80, Kluwer International Series in Engineering and Computer Science, Kluwer Academic Publisher, 1996.
3. Pentland, A., Essa, I., Darrell, T., Azarbayejani, A., and Sclaroff, S., "Visually Guided Animation," chapter in *Interactive Computer Animation,* N. Thalmann and D. Thalmann, Eds., pp. 143--164, Prentice Hall, 1996.
4. Sclaroff, S., Tacyher, L., La Cascia, M., "ImageRover: A Content-Based Image Browser for the World Wide Web," *Proc. IEEE Workshop on Content-Based Retrieval in Image Databases*, pp. 2--9, Jun., 1997.
5. Taycher, L., La Cascia, M., and Sclaroff, S., "Image Digestion and Relevance Feedback in the ImageRover WWW Search Engine," *Proc. International Conf. on Visual Information,* pp. 85--92, Dec., 1997.
6. Martin, J., Pentland, A., Sclaroff, S., and Kikinis, R., "Characterization of Neuropathological Shape Deformations," *IEEE Trans. Pattern Analysis and Mach. Intelligence* (PAMI), 30(2):97--112, Feb., 1998.
7. La Cascia, M., Sethi, S., and Sclaroff, S., "Combining Textual and Visual Cues for Content-Based Image Retrieval on the World Wide Web," *Proc. IEEE Workshop on Content-Based Access of Image and Video Libraries,* Jun., 1998.

8. Zhang, W., Dickinson, S., Sclaroff, S., Feldman, J., and Dunn, S., "Shape-Based Indexing for Content-Based Medical Image Retrieval," *Proc. IEEE Workshop on Biomedical Image Analysis,* Jun., 1998.
9. Rosales, R., and Sclaroff, S., "Improved Tracking of Multiple Humans with Trajectory Predcition and Occlusion Modeling," *Proc. IEEE Workshop on the Interpretation of Visual Motion,* Jun., 1998.

6. Invited lectures, panels, seminars, talks

Aug 1998 -- Aug 1999

1. Electrical and Computer Engineering Department, Boston University: *ImageRover: Content-Based Image Retrieval for the World Wide Web*, Boston, MA, Nov., 1998.
2. Harris Corporation: *Active Blobs*, Melbourne, FL, Nov., 1998.
3. University of Marlyand: *Active Blobs*, College Park, MD, Nov., 1998.
4. Boston College: *Active Blobs*, Boston, MA, Nov., 1998.
5. First International Workshop on Multimedia Internet Information Retrieval: *Content Based Image Browsing for the WWW*, Galassia Gutenberg Conference, Naples, Italy, Feb., 1999.
6. Boston University: *Mix and Match Features: Relevance Feedback and Indexing Strategies Employed in the ImageRover WWW Search Engine*,, Speech Processing Seminar Series, Dept. of Electrical and Computer Engineering, Boston University, Mar., 1999.
7. Yale University: *Active Blobs for Nonrigid Motion Tracking*, CVC group Vision Lunch seminar, Depts. of Electrical Engineering and Computer Science, Yale, Apr., 1999.

June, 1996 -- Aug 1998

1. Interval Research Corporation: *Deformable Shape Prototypes For Interactive Image Database Search,* Palo Alto, CA, June 1996.
2. University of Marlyand: ImageRover: *Content-Based Image Browser for the World Wide Web,* College Park, MD, December 1996.
3. IEEE Nonrigid and Articulated Motion Workshop: panel on future research directions, Puerto Rico, June 1997.
4. IEEE Workshop on Generic Object Recognition: *Generic Object Recognition with Active Blobs*, Puerto Rico, June 1997.
5. The Rutgers University Series on Human and Computer Vision, Rutgers University: *Active Blobs*, New Brunswick, NJ, August 1997.
6. IBM Almaden Research Center: *ImageRover,* Almaden, CA, December 1997.
7. Interval Research Corporation: *Active Blobs,* Palo Alto, CA, December, 1997.
8. Xerox Palo Alto Research Center: *ImageRover,* Palo Alto, CA, December, 1997.
9. Multi-Dimensional Signal Processing Laboratory Seminar, Boston University: *Active Blobs,* Boston, MA, March, 1998.
10. Microsoft Research Laboratory: *Active Blobs*, Redmond, WA, April, 1998.
11. Cambridge Research Lab, Digital Equipment Corporation: *Active Blobs,* Cambridge, MA, April, 1998.
12. Cambridge Research Lab, Digital Equipment Corporation: *ImageRover,* Cambridge, MA, April, 1998.
13. Institute for Robotics and Intelligent Systems, University of Southern California: *Active Blobs,* Los Angeles, CA, June 1998.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	Aug 1, 1999	Final Progress Report 5/96--8/99

4. TITLE AND SUBTITLE

Shape and Motion Categorization for Content Based Image and Video Database Search: Final Progress Report

5. FUNDING NUMBERS

G N00014-96-1-0661

6. AUTHOR(S)

Stan Sclaroff

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Computer Science Department
Boston University
111 Cummington Street
Boston, MA 02215

8. PERFORMING ORGANIZATION REPORT NUMBER

sclaroff-ONR-20

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Department of the Navy
Office of Naval Research
Ballston Centre Tower One
800 North Quincy Street
Arlington, VA 22217-5660

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The aim of this project is to represent shape categories for interactive, image database search. The work involves developing representations of images of objects to facilitate content-based retrieval and indexing, especially of deformable objects.

This is a final progress report for work conducted under a three year period, under the Office of Naval Research Young Investigator's Program.

14. SUBJECT TERMS

Image and video database retrieval, content-based indexing, nonrigid shape models

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
unclassified	unclassified	unclassified	UL

ENCL (1)